

# Making Minimal Solvers for Absolute Pose Estimation Compact and Robust

Viktor Larsson  
Lund University  
Lund, Sweden

viktorl@maths.lth.se

Zuzana Kukelova  
Czech Technical University  
Prague, Czech Republic

kukelzuz@fel.cvut.cz

Yinqiang Zheng  
National Institute of Informatics  
Tokyo, Japan

yqzheng@nii.ac.jp

## Abstract

*In this paper we present new techniques for constructing compact and robust minimal solvers for absolute pose estimation. We focus on the P4Pfr problem, but the methods we propose are applicable to a more general setting.*

*Previous approaches to P4Pfr suffer from artificial degeneracies which come from their formulation and not the geometry of the original problem. In this paper we show how to avoid these false degeneracies to create more robust solvers. Combined with recently published techniques for Gröbner basis solvers we are also able to construct solvers which are significantly smaller. We evaluate our solvers on both real and synthetic data, and show improved performance compared to competing solvers.*

*Finally we show that our techniques can be directly applied to the P3.5Pfr problem to get a non-degenerate solver, which is competitive with the current state-of-the-art.*

## 1. Introduction

Estimating the pose of a camera from minimal 2D-3D point correspondences is an important problem in geometric computer vision, as the minimal solvers often form the building blocks for 3D reconstruction frameworks. For estimating a calibrated camera only three points are required [11, 15]. If the intrinsic parameters are only partially known, more point correspondences are necessary. The most common situation in practice is that all intrinsic parameters except the focal length are known and the camera has some radial distortion. For the setting with unknown focal length and known or zero radial distortion, there have been many proposed minimal solvers using four points [23, 4, 25] and recently Wu [24] presented the first truly minimal solver using 3.5 points (by ignoring one image coordinate).

The absolute pose problem with unknown radial distortion and unknown focal length becomes minimal with four points and is usually called *P4Pfr*. It was first solved by Josephson and Byröd [14], however the elimination tem-

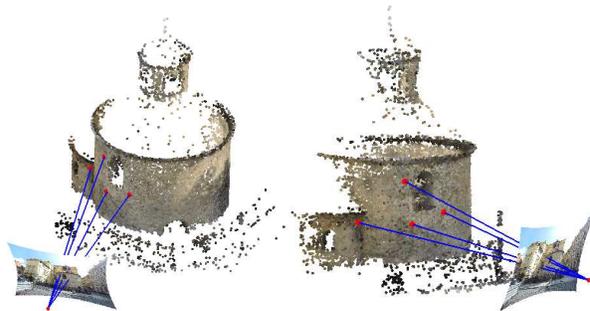


Figure 1. The camera pose, focal length and radial distortion are estimated from four 2D-3D correspondences. The examples above are from the *Rotunda* dataset (Section 4.4).

plate for this solver was quite large ( $1134 \times 720$ ), which limits its practical use. In [5] Bujnak *et al.* presented polynomial solvers with much smaller elimination templates by considering the planar and non-planar problems separately.

In [17] Kukelova *et al.* presented a non-minimal solver which uses five point correspondences. The solver is very fast but has the drawbacks of requiring more data and solving an overconstrained problem.

The P4Pfr problem becomes degenerate when the 3D points lie in a plane parallel to the image plane. In this case there exist infinitely many solutions by translating the camera towards the plane and changing the focal length. This degeneracy is inherent to the problem itself and cannot be avoided. However the previous solvers from [14] and [5] both suffer from additional degeneracies which are artificial in the sense that they come from the specific formulation used, instead of the geometry of the original problem.

In this paper we show how to avoid these unnecessary degeneracies and construct minimal solvers which are both more robust and have better performance. We focus on the problem of pose estimation in the case of unknown focal length and radial distortion. However the methods we present are not specific to this case and they are applicable to a more general setting.

The main contributions of this paper are

- We present new polynomial solvers for P4Pfr which

outperform the current state-of-the-art.

- Using clever parameterizations we avoid the nullspace and rotation degeneracies which are present in competing methods.
- Using the recently proposed elimination ideal technique [21] we are able to find significantly smaller elimination templates.
- Using a new template merging method we construct a solver that works for both planar and non-planar data.
- We also apply our approach to the P3.5Pf problem and get results comparable to the current state-of-the-art.

## 2. Background and Previous Work

### 2.1. Unknown Radial Distortion and Focal Length

If the image has undergone radial distortion, the standard pinhole camera model is no longer valid, and more complicated models are required. To handle this, different models have been proposed, *e.g.* [9, 10, 7]. One of the more popular models ([2, 13, 14, 5, 18]) is the one parameter division model presented by Fitzgibbon [10], since it provides a good trade-off between representation accuracy and model complexity.

The model assumes the undistorted image coordinates  $(u_u, v_u)$  are given by

$$(u_u, v_u) = \frac{1}{1 + k(u_d^2 + v_d^2)}(u_d, v_d) \quad (1)$$

where  $(u_d, v_d)$  are the distorted image coordinates observed in the image. The strength of the distortion is controlled by the parameter  $k$ , which is typically negative in the presence of barrel distortion.

In this model the projection equations can be written as

$$\lambda_i \begin{pmatrix} u_i \\ v_i \\ 1 + kd_i \end{pmatrix} = P X_i = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} \quad (2)$$

where  $d_i = u_i^2 + v_i^2$ . We assume that the camera has zero skew, unit aspect ratio and that the principal point lies in the center of the image. The camera matrix must then satisfy

$$P \sim K[R \ t], \quad K = \text{diag}(f, f, 1), \quad R^T R = I \quad (3)$$

Since the camera matrix has 7 degrees of freedom and the radial distortion parameter  $k$  is unknown, the problem becomes minimal with four 2D-3D point correspondences.

### 2.2. Minimal Solver from Josephson and Byröd

The first solution to the P4Pfr problem was presented in [14] by Josephson and Byröd. In the paper they proposed to parametrize the problem directly using quaternions for the

rotation matrix. The scale of the camera matrix was fixed by setting the first element of the quaternion to one. Finally, by using a clever choice of coordinate systems, they were able to eliminate the translation. This formulation leads to 5 equations in the 5 remaining unknowns (three quaternion parameters, the inverse focal length and the distortion parameter).

Using the techniques from [6] they constructed a Gröbner basis based solver for this system. The solver performs linear elimination on a matrix of size  $1134 \times 720$  followed by eigendecomposition of a  $24 \times 24$  matrix. The large template results in poor numerical stability and long running time, both of which make the solver unsuitable for practical applications.

The solver works for both planar and non-planar 3D points, but has a non-trivial degeneracy for any  $180^\circ$  rotation introduced by setting the first quaternion element to one. In addition, the number of solutions from this solver (24) is unnecessarily doubled, because of a hidden symmetry between the focal length and the rotation [24].

### 2.3. Minimal Solver from Bujnak et al.

In [5] Bujnak *et al.* presented another minimal solver for the P4Pfr problem. Their solver is based on the observation that if we eliminate the depth  $\lambda_i$  from the first two equations in (2), we get

$$v_i P_1 X_i - u_i P_2 X_i = 0, \quad i = 1, 2, 3, 4 \quad (4)$$

which is linear in the first two rows of the camera matrix and does not contain the radial distortion parameter  $k$ . Since there are four such equations, this can be rewritten as

$$M \mathbf{v} = \mathbf{0}, \quad M \in \mathbb{R}^{4 \times 8} \quad (5)$$

where  $\mathbf{v} = [p_{11}, p_{12}, p_{13}, p_{14}, p_{21}, p_{22}, p_{23}, p_{24}]$ . This is used to parametrize the first two rows of the camera matrix using only four parameters,

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 + \alpha_4 \mathbf{v}_4, \quad (6)$$

where  $\{\mathbf{v}_i\}_{i=1}^4 \subset \mathbb{R}^8$  is a basis for the nullspace of  $M$ . Since the reprojection equations are homogeneous in the camera matrix, the scale is fixed by setting  $\alpha_4 = 1$ . From (2) the remaining linearly independent equations can be written as

$$(1 + kd_i)P_1 X_i - u_i P_3 X_i = 0, \quad i = 1, 2, 3, 4 \quad (7)$$

Collecting the terms properly will lead to

$$A[p_{31}, p_{32}, p_{33}, p_{34}]^T = B[\alpha, k\alpha, k, 1]^T \quad (8)$$

where  $A \in \mathbb{R}^{4 \times 4}$ ,  $B \in \mathbb{R}^{4 \times 8}$  and  $\alpha = [\alpha_1, \alpha_2, \alpha_3]^T$ . Multiplying with the inverse of  $A$ , the third camera row is expressed in the four unknowns  $\alpha_1, \alpha_2, \alpha_3$  and  $k$ .

The left-most  $3 \times 3$  part of the camera matrix should correspond to  $KR$  (3). This gives constraints that the rows should be pairwise orthogonal and that the first two rows should have the same norm, as follows

$$p_{21}p_{31} + p_{22}p_{32} + p_{23}p_{33} = 0 \quad (9)$$

$$p_{11}p_{31} + p_{12}p_{32} + p_{13}p_{33} = 0 \quad (10)$$

$$p_{11}p_{21} + p_{12}p_{22} + p_{13}p_{23} = 0 \quad (11)$$

$$p_{11}^2 + p_{12}^2 + p_{13}^2 - p_{21}^2 - p_{22}^2 - p_{23}^2 = 0 \quad (12)$$

Using these equations Bujnak *et al.* [5] created a solver which performs Gaussian elimination on a template of size  $136 \times 152$ . Once the camera matrices are found, the focal length can be recovered by solving a quadratic polynomial. This formulation has 16 solutions, but only 12 of these are geometrically valid for the original problem.

This solver greatly reduces the template size compared to the solver from Josephson and Byröd [14]. However, it only works for non-planar 3D points. In the paper, the authors proposed a special solver to handle the planar case separately. The planar solver has an elimination template of size  $12 \times 18$ .

### 3. Our Approach for P4Pfr

Now we will present our approach for solving the P4Pfr problem. It builds on the formulation from Bujnak *et al.* [5], but improves it in three key aspects:

- The artificial degeneracy introduced by fixing the scale with  $\alpha_4 = 1$  is removed.
- Using the recent elimination ideal technique [21] we get a significantly smaller elimination template.
- We remove the planar degeneracy and create a unified solver that works for both planar and non-planar scenes.

#### 3.1. Removing Nullspace Degeneracy

In the solver by Bujnak *et al.* [5], the scale is fixed by setting  $\alpha_4 = 1$  in (6). This has the benefit of reducing the number of unknowns by one. However it introduces a degeneracy for any camera matrix which correspond to  $\alpha_4 = 0$ . Since the nullspace is only determined up to a  $4 \times 4$  change of variables, this essentially excludes a random set of camera matrices. It is unlikely in practice that the true solution has  $\alpha_4 = 0$  exactly, however any solutions close to these degenerate configurations can result in bad numerics (see Section 4.3 for experiments on this).

To avoid this degeneracy we propose a simple method for ensuring that the camera matrices which are excluded are geometrically uninteresting. To accomplish this we instead

fix the scale by setting the first depth to one, *i.e.*  $\lambda_1 = 1$ . Then for the first point the projection equations become

$$(u_1, v_1, 1 + kd_1)^T = PX_1 \quad (13)$$

Using this technique the first point now gives two linear constraints on the first two camera rows,  $P_1$  and  $P_2$ . The homogeneous linear system in (5) now becomes inhomogeneous,

$$M\mathbf{v} = \mathbf{b}, \quad M \in \mathbb{R}^{5 \times 8}, \quad \mathbf{b} \in \mathbb{R}^8 \quad (14)$$

and has one additional row. The solutions to this system can be parametrized as

$$\mathbf{v} = \mathbf{v}_0 + \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 \quad (15)$$

where  $M\mathbf{v}_0 = \mathbf{b}$  and  $\{\mathbf{v}_i\}_{i=1}^3 \subset \mathbb{R}^8$  forms a basis for the nullspace of  $M$ .

Note that this is essentially the same parametrization as before, but we have made sure that the degeneracy now instead occurs when the first point has zero depth, *i.e.* the first 3D point coincides with the camera center, an unrealistic scenario that never occurs in practice.

#### 3.2. New Camera Matrix Constraints

In [21] Kukulova *et al.* presented a new technique for using elimination ideals to construct smaller polynomial solvers. The approach is based on the observation that for many problems the equations can be divided into two groups; linear equations which depend on the data and non-linear equations which are independent of the data. By computing elimination ideals [8] for the non-linear equations, it is possible to eliminate some of the unknowns before constructing the elimination template. For a more detailed description of the process see [21]. In [21] the elimination ideal technique was applied to three relative pose problems and one homography estimation problem. Here we show how to apply this technique to absolute pose problems.

In the P4Pfr problem the nonlinear equations are<sup>1</sup>

$$P = \text{diag}(f, f, 1) [R \ t], \quad R^T R = sI \quad (16)$$

Computing the elimination ideal which eliminates all unknowns except for  $P$  yields the following set of equations

$$p_{13}^2 p_{32} - p_{21}^2 p_{32} - p_{22}^2 p_{32} - p_{12} p_{13} p_{33} - p_{22} p_{23} p_{33} = 0 \quad (17)$$

$$p_{12} p_{13} p_{32} + p_{22} p_{23} p_{32} - p_{12}^2 p_{33} + p_{21}^2 p_{33} + p_{23}^2 p_{33} = 0 \quad (18)$$

$$p_{11} p_{13} p_{32} + p_{21} p_{23} p_{32} - p_{11} p_{12} p_{33} - p_{21} p_{22} p_{33} = 0 \quad (19)$$

$$p_{13}^2 p_{31} - p_{22}^2 p_{31} + p_{21} p_{22} p_{32} - p_{11} p_{13} p_{33} = 0 \quad (20)$$

$$p_{12} p_{13} p_{31} + p_{22} p_{23} p_{31} - p_{11} p_{12} p_{33} - p_{21} p_{22} p_{33} = 0 \quad (21)$$

in addition to the constraints (9)–(12). These constraints ensure that the first  $3 \times 3$  part of the camera matrix can be

<sup>1</sup>We add the unknown  $s$  since the camera matrix is only determined up to scale.

factorized as  $\text{diag}(f, f, 1)R$ , where  $R$  is a scaled rotation. To the best of our knowledge, the constraints (17)–(21) are new and have not been used in the computer vision literature before.

After adding these new equations, the formulation now correctly has 12 solutions, in contrast to 16 in [5] and 24 in [14]. Using the automatic solver generator from [16], we generate a polynomial solver with an elimination template of size  $28 \times 40$ . While the solver is significantly smaller than the non-planar solver from [5] ( $136 \times 152$ ), it also suffers from the same planar degeneracy.

In [21] the authors also considered radial distortion, but for relative pose. To avoid the extra non-linearity introduced by the radial distortion they used a lifting approach. We tried to apply the same technique for our problem and also eliminate the radial distortion, but the resulting solvers were too large for practical use ( $296 \times 330$ ).

### 3.3. Removing Planar Degeneracy

In this section we will extend the formulation to handle both planar and non-planar scenes.

For planar scenes we can without loss of generality assume that all  $z_i = 0$ , *i.e.* the 3D points lie in the  $xy$ -plane. As was noted in Bujnak *et al.* [5], for such scenes the third column of the  $A$  matrix in (8) becomes zero, and it is impossible to eliminate  $p_{33}$  this way. To avoid this situation we instead add  $p_{33}$  as an additional unknown. Then using only three of the four points we can express  $p_{31}$ ,  $p_{32}$  and  $p_{34}$  in the unknowns  $\alpha$ ,  $k$  and  $p_{33}$ , *i.e.*

$$A[p_{31}, p_{32}, p_{34}]^T = B[\alpha, k\alpha, k, p_{33}, 1]^T \quad (22)$$

where  $A \in \mathbb{R}^{3 \times 3}$  and  $B \in \mathbb{R}^{3 \times 9}$ . Note that this works for both planar and non-planar data. Since we have only used three of the four equations (7), we need add the last projection equation separately. Using the first point (which was used to fix the scale (13)) this equation is simply

$$1 + kd_1 = P_3 X_1 \quad (23)$$

Studying the equations in Macaulay2 [12] we found that this formulation also has 12 solutions. Using the automatic generator from [22] we generated a solver with template size  $38 \times 50$ . To our surprise this solver did not directly work for planar data.

Further investigations of the problem in Macaulay2 [12] revealed that for planar instances, the monomial basis,

$$\{1, \alpha_1, \alpha_2, \alpha_3, k, p_{33}, \alpha_1\alpha_3, \alpha_2^2, \alpha_2\alpha_3, \alpha_3^2, \alpha_3k, \alpha_3p_{33}\} \quad (24)$$

which was used for the quotient space  $\mathbb{C}[X]/I$  became linearly dependent (in the quotient space). Thus making it impossible to express the action matrix using it. Furthermore we found that if we added the monomial  $\alpha_1\alpha_2k$  to the basis,

it would span the quotient space for both planar and non-planar data. Since the structure of the ideal is different for the two types of instances, different monomials are needed in the elimination template. Using the automatic generator [22] we created elimination templates for both planar and non-planar instances and then constructed a single template by taking the union of all the necessary equations. Since the ideals are very similar for both cases, we were able to find a merged template which is only slightly bigger than the template created using non-planar data only.

The final solver performs gaussian elimination on a single template of size  $40 \times 50$  and then solves a  $13 \times 13$  eigenvalue problem (due to the extra basis element). However for any instance only 12 of the eigenvectors correspond to actual solutions. This solver does not suffer from any of the artificial degeneracies present in previous solvers and it works for both planar and non-planar data. Compared to the current state-of-the-art general solver from Josephson and Byröd [14] the size is orders of magnitude smaller ( $40 \times 50$  vs.  $1134 \times 720$ ).

## 4. Experimental Evaluation

### 4.1. Numerical Stability

To evaluate the stability and accuracy of the new P4Pfr solvers we use a similar experiment setup as was used in [25, 24]. We generate synthetic scenes by uniformly sampling four 3D points in the box  $[-2, 2] \times [-2, 2] \times [2, 8]$  in the camera’s local coordinate system. The 3D points are then transformed by a random rotation and translation. The focal length is randomly chosen in the interval  $f_{gt} \in [0.5, 2.5]$ . Radial distortion using the division model [10] was added to all image points to generate noiseless distorted points. The radial distortion parameter was randomly drawn from the interval  $k_{gt} \in [-0.45, 0]$ .

To perform the experiment we generated 10000 random scenes as described above. Figure 2 shows the histograms of the  $\log_{10}$  relative errors in the estimated focal length and radial distortion parameter obtained by selecting the real root closest to the ground truth values  $f_{gt}$  and  $k_{gt}$ . We also ran the same experiment for planar scenes, generated by projecting the four 3D points to the closest plane using SVD. The results are shown in Figure 3.

We can see that the new general P4Pfr ( $40 \times 50$ ) solver (blue) is stable for both the planar and non-planar setting. Moreover, the new solver is more stable than the state-of-the-art general solver from Josephson and Byröd [14] (green) for non-planar setting. The same holds true for the new non-planar solver (red) compared with the state-of-the-art non-planar solver from Bujnak *et al.* [5] (cyan).

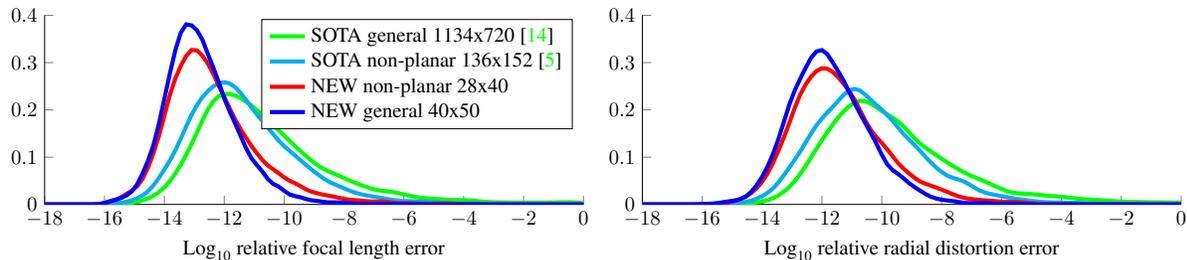


Figure 2. Histograms of  $\log_{10}$  relative errors of the estimated focal lengths (left) and radial distortions (right) for non-planar scenes.

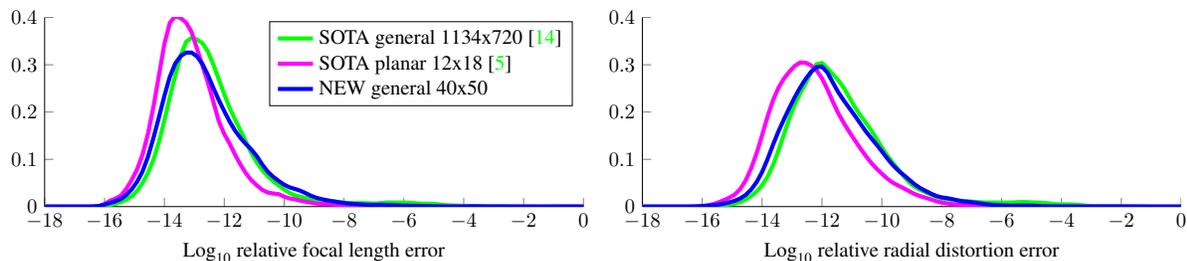


Figure 3. Histograms of  $\log_{10}$  relative errors of the estimated focal lengths (left) and radial distortions (right) for planar scenes.

## 4.2. Noise Experiment

In the next experiment we studied the performance of the new solvers in the presence of image noise. We again compare both presented solvers (the general solver (blue) and the non-planar solver (red)) with the the state-of-the-art general solver from Josephson and Byröd [14] (green) and the state-of-the-art non-planar solver from Bujnak *et al.* [5] (cyan).

In this experiment we used the same setup as was used in the previous stability experiment, however with the fixed ground truth focal length  $f_{gt} = 1.5$  and the fixed radial distortion  $k_{gt} = -0.4$ . For each noise level 1000 estimates for random scenes and camera positions were made. Figure 4 (Left) shows the median focal length errors for different noise levels. Figure 4 (Right) shows the errors for the focal lengths using the Matlab boxplot function which shows values 25% to 75% quantile as a box with horizontal line at median. The crosses show data beyond 1.5 times the interquartile range. The errors for the radial distortion parameter can be found in the supplementary material.

All minimal solvers perform equally well. This is caused by the fact that these solvers are all algebraically equivalent. The only difference is caused by numerical instabilities. This is *e.g.* visible in Figure 4 (Right) by the green crosses for the state-of-the-art solver by Josephson and Byröd [14] in the noiseless case.

### 4.2.1 Estimation from Non-Minimal Point Sets

In [17] Kukulova *et al.* presented a non-minimal solver which uses five point correspondences. To perform a com-

parison with this solver we generated a fifth point for each of the scenes. The results are included in Figure 4 and we can see that in the presence of noise the performance is superior compared to the four point solvers, which is reasonable since more data is used.

Since our solver is based on a nullspace parametrization it is possible to use five points in our solver as well. For estimation with non-minimal point sets the only changes we need to make is to compute an approximate nullspace in (14) using SVD and to solve (22) in a least squares sense. In Figure 4 we can see that our solver using 5 points (magenta) is more accurate for noisy data compared to the solver from [17] (yellow).

### 4.3. Stability Close to Degenerate Configurations

In this section we evaluate the stability of the polynomial solvers close the degenerate configurations. First we consider the quaternion based degeneracy in the solver from Josephson and Byröd [14]. We generated random scenes where the ground truth camera pose was close to the degenerate configuration (i.e. first element of the quaternion representing the rotation is close to zero). Figure 5 shows the median relative error in the focal length as the first element of the quaternion tends to zero.

Next we consider the nullspace based degeneracy from Section 3.1. To perform the experiment we randomly generate a scene and computed a basis for the nullspace (as in (5)). We then find the coefficients  $\alpha$  which correspond to the ground truth camera matrix. Next we perform a random rotation in the nullspace which brings  $\alpha_4$  close to zero. Figure 6 shows the result for our solver both using the

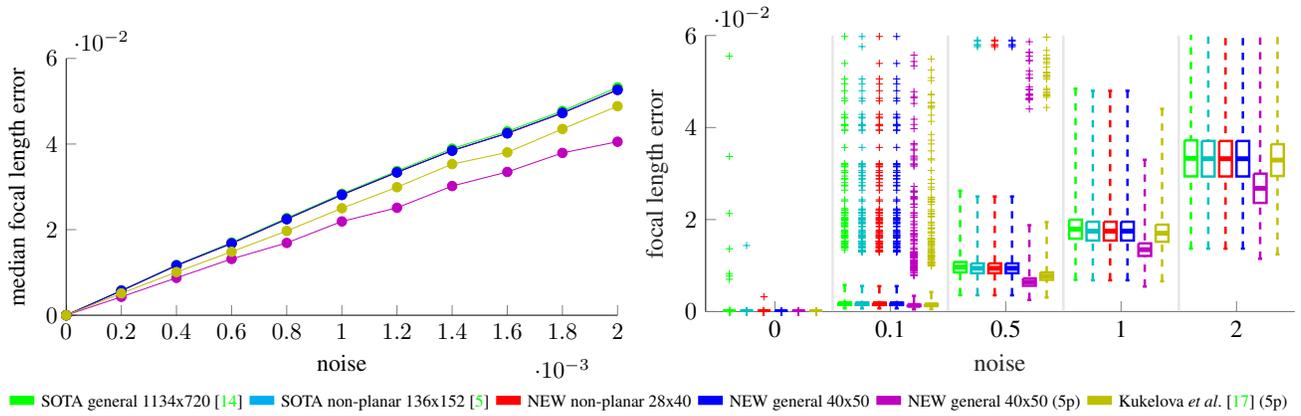


Figure 4. Comparison of the errors in the focal length estimated by different solvers for varying levels of noise. The ground truth values were set to  $f_{gt} = 1.5$  and  $k_{gt} = -0.4$ . *Left*: Median focal length error. *Right*: Boxplot of focal length errors.

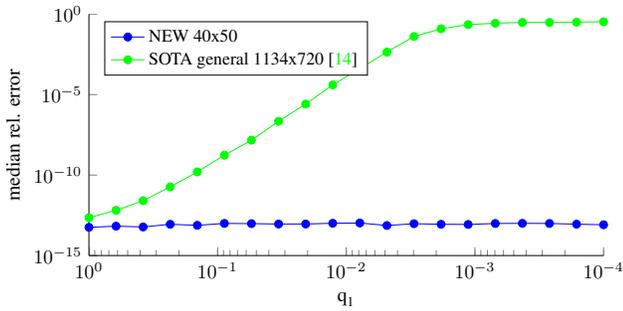


Figure 5. The relative focal length error as the first element of the quaternion approaches zero. Each point shows the median error over 1000 instances.

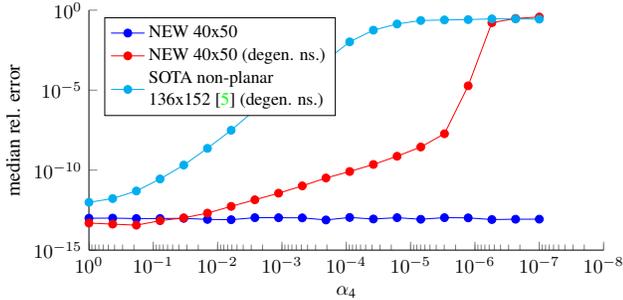


Figure 6. The relative focal length error as the last nullspace coefficient  $\alpha_4$  approaches zero. Each point shows the median error over 1000 instances. For our solver we show the results both with and without the scale fixing in Section 3.1.

parametrization from Section 3.1 and using the degenerate nullspace created as above. Since this degeneracy is also present in the solver from Bujnak *et al.* [5] we include the results of running their solver with the degenerate nullspace basis as well.

Finally we consider close-to-planar degeneracy. To evaluate the performance of our new P4Pfr solvers on close-to-planar scenes we use a similar experiment setup as was

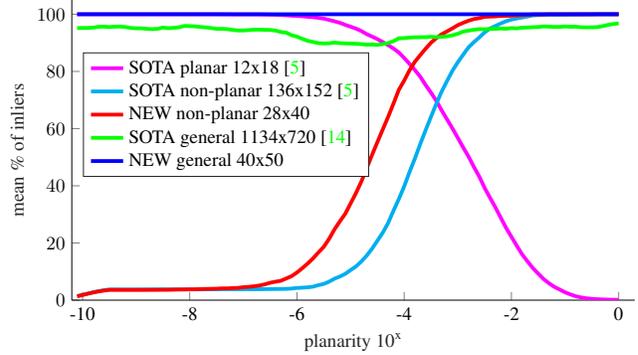


Figure 7. Mean of the number of inliers for a near planar scene.

used in [5]. We created a synthetic scene where we were able to control the scene planarity by a scalar value  $a$ . First, we generated three random non-collinear 3D points. These three points define our plane. Then we randomly generated the fourth point at the distance  $sa$  from the plane, where the scale  $s$  was the maximum distance from the first three points to their center of gravity. The fourth point was generated such that its distance from the center of gravity was not greater than  $s$ . This means that for  $a = 0$  we got four points on the plane and for  $a = 1$  we got a well-defined non-planar four-tuple of 3D points. For each given planarity value  $a$  we created a scene consisting of these four 3D points and an additional 100 random 3D points. Each 3D point was projected by a camera with random feasible orientation and position,  $f_{gt} = 1.5$  and  $k_{gt} = -0.4$ . The first four points were kept noise-free, while we added some small noise (standard deviation of 0.5 pixels) to the remaining 100 points.

For each given planarity value  $a$  we computed the camera pose, focal length, and radial distortion from the first four-tuple of correspondences. This four-tuple was not affected by a noise and hence the only deviation from the ground truth solutions comes from the numerical instabil-

ity of the solvers itself. To evaluate the impact of this instability we used the estimated camera pose, focal length and radial distortion to project all remaining 3D points to the image plane. Then we measured the number of inliers, i.e. the number of points that were projected closer than one pixel to its corresponding 2D image point.

In this experiment, for each given planarity value  $a$  we created 100 random scenes. Figure 7 shows the mean number of inliers found by different solvers. It can be seen that the new non-planar solver (red) performs better than the state-of-the-art non-planar solver from Bujnak *et al.* [5] (cyan) for close-to-planar scenes and is therefore more suitable for a “joined general solver” presented in [5]. The new general P4Pfr solver (blue) doesn’t have problems with close-to-planar or planar scenes and this solver was able to find all inliers. The state-of-the-art general solver [14] (green) has some problems with numerical stability due to the decomposition of a huge template matrix ( $1134 \times 720$ ). Therefore, the average number of inliers returned by this solver was less than 100.

#### 4.4. Evaluation on Real Images

Finally we evaluate our method on real image data. We consider the *Rotunda* dataset [19] and the “planar” *Graffiti* dataset [20]. The *Rotunda* dataset consists of 62 images captured using a GoPro Hero4 camera with significant radial distortion. The *Graffiti* dataset consists of 12 images captured using GoPro Hero3 camera and 7 images captured using a HTC Desire 500 mobile phone. Some example images are shown in Figure 8. Using the RealityCapture software [1] we built a 3D reconstructions of both scenes. The *Rotunda* reconstruction contains 170994 3D points and the average reprojection error was 1.4694 pixels over 549478 image points. The *Graffiti* reconstruction contains 26078 3D points and the average reprojection error was 1.0778 pixels over 91518 image points.

Then to perform the experiment we used the 3D model to estimate the pose of each image using the new minimal solver ( $40 \times 50$ ) in a RANSAC framework. Since the dataset only contains image data, we used the camera and distortion parameters obtained from RealityCapture as ground truth for the experiment. Table 1 shows the errors for the focal length and radial distortion, as well as the camera pose. Overall the errors are quite small, with slightly larger errors for the more difficult planar dataset (*Graffiti*).

### 5. Our Approach for P3.5Pf

Now we show how to apply our approach to the closely related problem of pose estimation with unknown focal length. This problem has 7 degrees of freedom and is over-constrained with four points. In [24] Wu presented a minimal solver using 3.5 points (*i.e.* ignoring one of the image coordinates for one of the points). This solver works for

Dataset	<i>Rotunda</i>			<i>Graffiti</i>		
	avg.	med.	max	avg.	med.	max
Focal (%)	0.08	0.07	0.29	0.44	0.33	1.76
Distortion (%)	0.51	0.45	1.85	2.16	1.23	8.85
Rotation (degree)	0.03	0.03	0.10	0.12	0.11	0.27
Translation (%)	0.07	0.07	0.26	1.30	0.70	5.06

Table 1. Errors for the real datasets. The errors are relative to the ground truth for all except rotation which is shown in degrees.

both planar and non-planar data but has a degeneracy introduced by setting one quaternion element to one.

In this section we develop a new solver for this problem which has comparable performance to [24], but does not introduce any artificial degeneracies. The approach is essentially the same as for our P4Pfr solver, but the formulation is simplified slightly since the projection equations (2) become completely linear when the radial distortion is removed. Each point correspondence now gives us two linearly independent equations in the camera matrix,

$$P_1 X_i - u_i P_3 X_i = 0, \quad P_2 X_i - v_i P_3 X_i = 0 \quad (25)$$

for  $i = 1, 2, 3, 4$ . Ignoring one of these eight equations gives a minimal problem. Using the same trick as in Section 3.1, we fix the scale by setting the first depth to one,

$$u_1 = P_1 X_1, \quad v_1 = P_2 X_1, \quad 1 = P_3 X_1 \quad (26)$$

Rewriting the linear constraints as

$$M \mathbf{v} = \mathbf{b}, \quad M \in \mathbb{R}^{8 \times 12}, \quad (27)$$

we can parametrize the problem with only four unknowns using the nullspace of  $M$ . Note that for this problem the constraints on the camera matrix are exactly same as in Section 3.2. Using these 9 equations in the four unknowns, we generated a polynomial solver using the automatic generator from [16]. The resulting solver has a template of size  $25 \times 35$ , comparable to the current state-of-the-art [24] ( $20 \times 30$ ). However in contrast to [24] this formulation does not contain any additional degeneracies.

In this formulation the problem has 10 solutions for general data, and 8 solutions for planar data. For this problem the quotient basis and template we found from the non-planar instances works for planar instances as well. In the case of planar data the solver still returns 10 solutions, but only 8 will correspond to actual solutions.

#### 5.1. Experiment

To evaluate the stability and accuracy of the polynomial solver we use a similar experiment setup as was used in [25, 24]. We generated random instances by uniformly sampling four 3D points in the box  $[-2, 2] \times [-2, 2] \times [2, 8]$

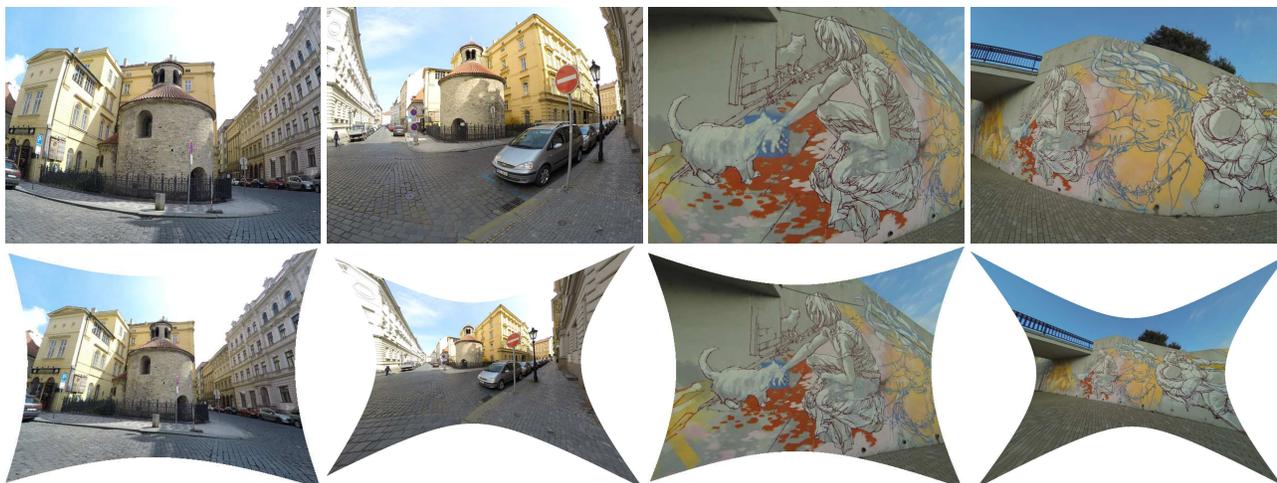


Figure 8. Some examples of the images in the *Rotunda* and *Graffiti* datasets. *Top*: Original images. *Bottom*: Undistorted images.

in the camera’s local coordinate system. The points were transformed by a random rotation and translation. The focal length was chosen uniformly in the interval  $[200, 2000]$ . Figure 9 shows the  $\log_{10}$  relative focal length error for 1000 instances for both planar and non-planar data. For comparison we include the results for the best ratio and distance based solvers from Bujnak [3] and the GP4Pf solver from Zheng *et al.* [25]. We can see that the new P3.5Pf solver (blue) is stable for both the planar and non-planar setting. We were unable to directly compare with the solver from [24], since the code has not been made available. However in [24] they report comparable results to the solver from Zheng *et al.* [25].

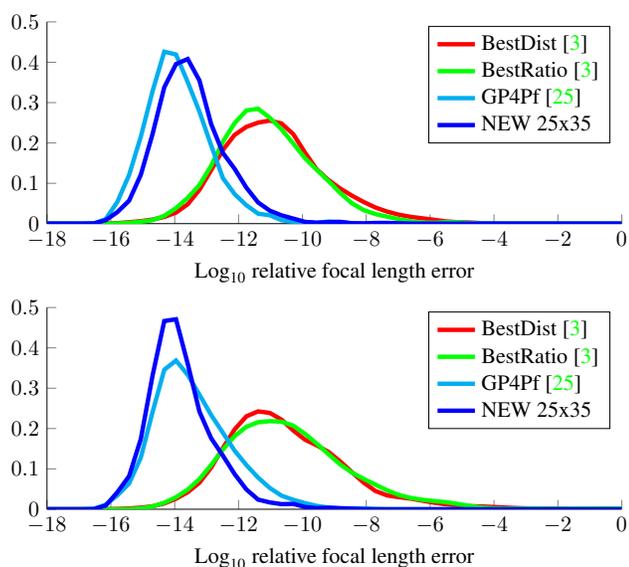


Figure 9. Stability of the P3.5Pf solver. The figure shows the relative focal length errors over 1000 random instances. *Top*: General 3D data. *Bottom*: Planar 3D data.

## 6. Conclusions

In this work we have presented new tricks for minimal absolute pose solvers. We have applied them to two problems, P4Pf and P3.5Pf, but we believe that they could be applicable to more problems.

First we removed the degeneracy introduced by fixing the scale in the nullspace parametrization by manually choosing which camera matrices to exclude. It is possible that this trick could be applied for other problems as well, *e.g.* in relative pose problems where nullspace parametrizations are commonly used.

Next we applied the elimination ideal techniques from [21] to get new constraints on the camera matrix. These constraints are satisfied whenever the focal length is the only unknown intrinsic parameter. It is possible that this technique could be applied to settings with other partial calibrations as well.

Finally we were able to construct a solver which worked for both planar and non-planar scenes. The key idea was to make sure that the monomial basis spanned both quotient spaces and then creating a single merged elimination template which contains the equations necessary to solve for both types of instances. We believe that this template merging strategy could be applied to other problems as well. Not only for planar/non-planar degeneracies, but any time the structure of the problem depends on the input data.

## 7. Acknowledgements

This work was finished when Viktor Larsson and Zuzana Kukelova were visiting the National Institute of Informatics (NII), Japan, funded in part by the NII MOU/Non-MOU International Internship/Exchange Program. Zuzana Kukelova was supported by The Czech Science Foundation Project GACR P103/12/G084.

## References

- [1] RealityCapture. [www.capturingreality.com](http://www.capturingreality.com). 7
- [2] J. P. Barreto and K. Daniilidis. Fundamental matrix for cameras with radial distortion. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 1, pages 625–632. IEEE, 2005. 2
- [3] M. Bujnak. Algebraic solutions to absolute pose problems. Ph. D. dissertation. Czech Technical University, Prague., 2012. 8
- [4] M. Bujnak, Z. Kukelova, and T. Pajdla. A general solution to the p4p problem for camera with unknown focal length. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008. 1
- [5] M. Bujnak, Z. Kukelova, and T. Pajdla. New efficient solution to the absolute pose problem for camera with unknown focal length and radial distortion. In *Asian Conference on Computer Vision*, pages 11–24. Springer, 2010. 1, 2, 3, 4, 5, 6, 7
- [6] M. Byröd, K. Josephson, and K. Åström. A column-pivoting based strategy for monomial ordering in numerical gröbner basis calculations. In *European Conference on Computer Vision*, pages 130–143. Springer, 2008. 2
- [7] D. Claus and A. W. Fitzgibbon. A rational function lens distortion model for general cameras. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 213–219. IEEE, 2005. 2
- [8] D. Cox, J. Little, and D. O’shea. *Ideals, varieties, and algorithms*, volume 3. Springer, 1992. 3
- [9] F. Devernay and O. D. Faugeras. Automatic calibration and removal of distortion from scenes of structured environments. In *SPIE’s 1995 International Symposium on Optical Science, Engineering, and Instrumentation*, pages 62–72. International Society for Optics and Photonics, 1995. 2
- [10] A. W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–I. IEEE, 2001. 2, 4
- [11] X.-S. Gao, X.-R. Hou, J. Tang, and H.-F. Cheng. Complete solution classification for the perspective-three-point problem. *IEEE transactions on pattern analysis and machine intelligence*, 25(8):930–943, 2003. 1
- [12] D. R. Grayson and M. E. Stillman. Macaulay 2, a software system for research in algebraic geometry, 2002. 4
- [13] H. Jin. A three-point minimal solution for panoramic stitching with lens distortion. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008. 2
- [14] K. Josephson and M. Byröd. Pose estimation with radial distortion and unknown focal length. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 2419–2426. IEEE, 2009. 1, 2, 3, 4, 5, 6, 7
- [15] L. Kneip, D. Scaramuzza, and R. Siegwart. A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 2969–2976. IEEE, 2011. 1
- [16] Z. Kukelova, M. Bujnak, and T. Pajdla. Automatic generator of minimal problem solvers. In *European Conference on Computer Vision*, pages 302–315. Springer, 2008. 4, 7
- [17] Z. Kukelova, M. Bujnak, and T. Pajdla. Real-time solution to the absolute pose problem with unknown radial distortion and focal length. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2816–2823, 2013. 1, 5, 6
- [18] Z. Kukelova, M. Byröd, K. Josephson, T. Pajdla, and K. Åström. Fast and robust numerical solutions to minimal problems for cameras with radial distortion. *Computer Vision and Image Understanding*, 114(2):234–244, 2010. 2
- [19] Z. Kukelova, J. Heller, M. Bujnak, A. Fitzgibbon, and T. Pajdla. Efficient solution to the epipolar geometry for radially distorted cameras. In *Proceedings of the IEEE international conference on computer vision*, pages 2309–2317, 2015. 7
- [20] Z. Kukelova, J. Heller, M. Bujnak, and T. Pajdla. Radial distortion homography. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 639–647, 2015. 7
- [21] Z. Kukelova, J. Kileel, B. Sturmfels, and T. Pajdla. A clever elimination strategy for efficient minimal solvers. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017. 2, 3, 4, 8
- [22] V. Larsson, K. Åström, and M. Oskarsson. Efficient solvers for minimal problems by syzygy-based reduction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017. 4
- [23] B. Triggs. Camera pose and calibration from 4 or 5 known 3d points. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 1, pages 278–284. IEEE, 1999. 1
- [24] C. Wu. P3. 5p: Pose estimation with unknown focal length. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2440–2448, 2015. 1, 2, 4, 7, 8
- [25] Y. Zheng, S. Sugimoto, I. Sato, and M. Okutomi. A general and simple method for camera pose and focal length determination. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 430–437, 2014. 1, 4, 7, 8